Air & Water
Borne Diseases

Cooke III et al., Air Water Borne Diseases 2013, 2:2
http://dx.doi.org/10.4172/2167-7719.1000111

Research Article Open Access

Development of a Sampling Scheme for West Nile Virus Testing in Urban and Rural Environments

William H Cooke III[1]*, Katarzyna Grala[2], Christopher P Brooks[3] and Linda MW Pote[4]

[1]*GIS & Remote Sensing, Mississippi State University, USA*
[2]*GIS & Remote Sensing, Mississippi State University, USA*
[3]*Department of Biological Sciences, Mississippi State University, USA*
[4]*Department of Basic Sciences, College of Veterinary Medicine, Mississippi State University, USA*

Abstract

A sampling scheme has been developed for West Nile Virus (WNV) testing in urban and rural environments for four study sites proximal to cities in Mississippi that are representative of a wide variety of ecological conditions. Population clustering and landscape characteristics were taken into consideration for a development of sample stratification scheme that minimizes redundancy for testing for WNV in bird and mosquito populations. The stratification scheme divides census block groups into rural and urban polygons and characterizes similar landscape conditions associated within these polygons. Landform characteristics for water percentage and terrain roughness were measured for these polygons using a GIS zonal statistics function and classified into three categories of roughness and three categories of percent water. Combination of these categories resulted in nine unique index values that provide a simple measure of landscape characteristics. Locating similar index values for urban and rural polygons enables matching of sample sites with similar landscape properties for all four study sites. The implementation of this sample design for testing birds and mosquitoes is important for determining whether WNV disease repository originates in urban environments and spreads to rural areas or vice versa.

Keywords: GIS; Sampling; Stratification; West nile virus; Census; Landform; Landscape; Index

Introduction

West Nile Virus (WNV) is one of the most widespread arboviruses in the United States and it has been annually linked to numerous human mortalities since its introduction in New York in 1999 [1]. In Mississippi, first incidence of WNV was detected in 2001 and to this day the virus is considered a significant health hazard for the state residents [2]. In Mississippi, similar to other states, there are WNV "hot spots" that have occurred in more urbanized areas; however, there are also rural areas in the state that have had a higher WNV incidence per capita [3].

Numerous studies focused on the ecological factors involved in mosquito-borne pathogens occurring in humans in urban centers in the United States [4-6]. However, there is less information available on the ecological factors involved in transmission of these pathogens in humans and animals in non-urban settings [1,7]. This paper focuses on development of a sampling scheme that minimizes redundancy for testing for WNV in bird and mosquito populations in both urban and rural environments in four Mississippi cites that encompass a wide range of ecological, hydrologic, and social conditions across the state.

Methods

Justification for stratified sampling

The 'best' spatial sampling scheme depends on the spatial characteristics of variability in data associated with the phenomena under investigation. Randomly sampled data can exhibit a high degree of spatial autocorrelation when samples are drawn from nearby locations [8]. Studies associated with the geography of a disease will naturally tend to cluster in more densely populated areas resulting in a *'first order effect'* that often results in violation of the assumption of sample independence leading to data redundancy. Redundancy can be reduced through identification of spatial strata that minimize within-strata sample variance and maximize among-strata variance [9]. A stratification design is proposed that takes population clustering and landscape characteristics into consideration for development of sample strata to minimize redundancy for testing for WNV in bird and mosquito populations. The stratification design first divides census block groups by employing an urban/rural population continuum threshold and then characterizes similar landscape conditions associated within the urban and rural polygons. The result is that landscape conditions are held as constant as possible while the population per square mile is allowed to vary.

Development of the urban - rural continuum and choice of threshold using census block group data

The U.S. Census Bureau defines the threshold between urban and rural areas as population density value of 1000 persons per square mile [10]. This density threshold is also used to define the 'urban fringe'. Although there are a variety of additional urban/rural classification criteria, this 'urban fringe' threshold value is useful for geographic information system (GIS) analysis aimed at dividing census block groups into urban and rural polygons. In this proposed sampling scheme, the 'urban fringe' threshold was expanded to create a set of polygons that are on the urban/rural boundary and another set of mutually exclusive polygons that are clearly within either urban or rural areas.

***Corresponding author:** William H Cooke III, GIS and Remote Sensing, Mississippi State University, USA, E-mail: whc5@geosci.msstate.edu

Received June 24, 2013; **Accepted** August 09, 2013; **Published** August 16, 2013

Citation: Cooke III WH, Grala K, Brooks CP, Pote LMW (2013) Development of a Sampling Scheme for West Nile Virus Testing in Urban and Rural Environments. Air Water Borne Diseases 2: 111. doi: 10.4172/2167-7719.1000111

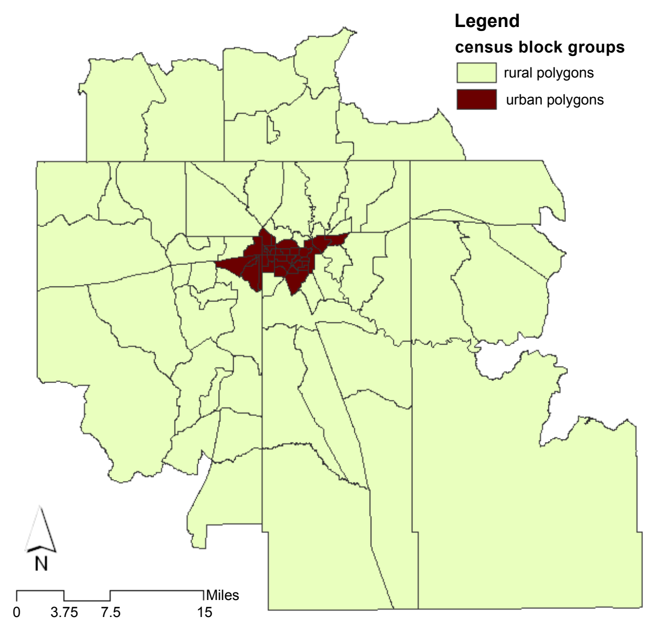

Figure 1: Urban and rural census block groups identified for Site 4.

The 'boundary' polygons were not included in the stratification design. For each census block group polygon marked as urban or rural, a GIS zonal analysis was performed for characterizing landscape conditions within each block group polygon. The 'zonal statistics' function simply summarizes information within each block group (zone) for any continuous variable mapped as a raster data type associated with landscape properties [11]. For this sampling scheme, topographic variation and percent water were standardized by census block group area and summarized by each census block group polygon. Figure 1 shows an example of urban and rural census block groups defined for Site 4.

Variable justification

The decision to use percent water per area and terrain roughness to describe potential for mosquito habitat was predicated on previous work by [2] where the overall median measure of infection risk by zip code indicated that 67% human cases occurred in the high-risk category. Weighting criteria for determination of risk categories were based on significant p-values from t-tests for road density (p-value=.001), stream density (p-value=.010) and slope (p-value=.028). Density analyses that were performed for the study in GIS were carried out using a kernel density approach that requires specification of search radii (bandwidth) that could potentially lead to over- or under-specification of density. One method for reducing the subjectivity associated with kernel density approaches for determining density is use of standardized measures for spatially extensive variables [12]. This was accomplished for development of the sampling protocol described in this paper by summarizing variable information by each zone. Population density per area in each zone was used to categorize the urban-rural continuum. Slope was significantly different for zip codes of WNV occurrence versus zip codes of non-occurrence in 2002 and 2003 in Mississippi. The decision to use the first derivative of slope (roughness) for development of this sampling scheme was based on the need to better describe the characteristics of surface water movement across landscapes. Modeling roughness helps in determination of surface water runoff characteristics [13], that are a determining factor in whether water can pond long enough to facilitate mosquito breeding. Measurements of stream density using the kernel density approach are subject to the same bandwidth problems described for road density and are further complicated by the fact that stream density can actually be high in areas of high landscape roughness with numerous intermittent streams or in areas with braided stream characteristics. It is the authors' opinion that the standardized ratio of water area per census zone coupled with landscape roughness provides the best interactive combination of variables that map the potential for persistent water on the landscape.

Landscape continuum characterization by topography and percent water per census block group using an Index Model

In order to limit variation of landscape properties within each census block group a GIS analysis was used to compare within- and among-strata landscape properties. For the purposes of this sampling scheme, comparisons of mean topographic variation (roughness) and mean percent water within each census block group were calculated. A measure of topographic variation can be obtained by calculating the second derivative of elevation which creates a terrain roughness map [14]. A GIS 'zonal statistics' function was used to summarize the mean roughness per census block group polygon, whereas the percent water per polygon was obtained by a zonal analysis of National Land Cover Dataset by summarizing water versus non-water land cover using 30-meter raster cells. The continuous values for mean roughness and percent water were binned into three categories of roughness and three categories of percent water using a quantile histogram division method. The quantile or equal count classification scheme simply assigns the same number of observations to each class [11]. Water classes were assigned index values of 1, 2, and 3 (1=small percentage of water, 2=medium percentage of water, 3=large percentage of water) and terrain roughness classes were assigned index values of 10, 20, and 30 (10=small roughness, 20=medium roughness, 30=large roughness). Summing the index values resulted in nine unique combinations of landscape properties that are defined by the following index values as (Figure 2):

11=water percent small, roughness small

12=water percent small, roughness medium

13=water percent small, roughness large

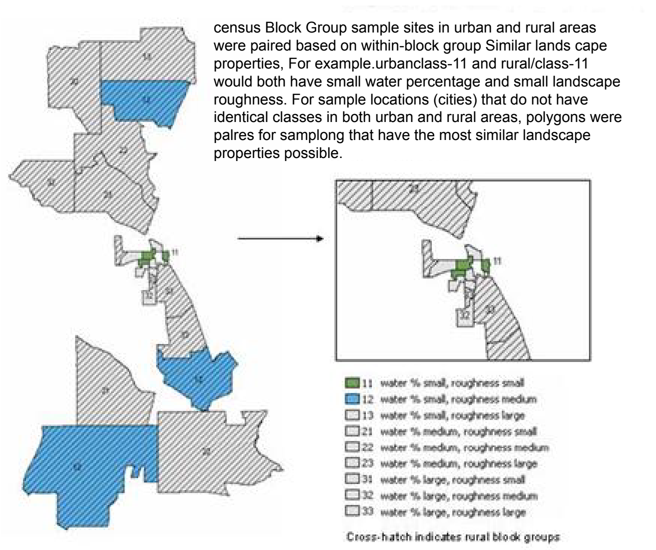
census Block Group sample sites in urban and rural areas were paired based on within-block group Similar lands cape properties, For example.urbanclass-11 and rural/class-11 would both have small water percentage and small landscape roughness. For sample locations (cities) that do not have identical classes in both urban and rural areas, polygons were palres for samplong that have the most similar landscape properties possible.

Figure 2: Site 1 Census block groups used for sample site selection.

	Site 1			Site 2			Site 3			Site 4		
Unique Feature Combinations	Exists	No.		Exists	No.		Exists	No.		Exists	No.	
		U	R		U	R		U	R		U	R
11	X	3	0	X	10	0	X	36	0	X	16	0
12	X	0	3	X	0	1	X	12	1	X	2	0
13	X	0	2	-	0	0	X	3	3	X	0	1
21	X	3	1	X	8	0	X	28	0	X	2	0
22	X	0	2	X	0	3	X	13	0	X	1	8
23	X	1	1	-	0	0	X	9	5	X	2	6
31	X	1	0	X	5	0	X	32	0	X	0	1
32	X	1	2	X	0	1	X	5	4	X	2	6
33	X	1	2	X	1	4	X	7	7	X	0	10

Table 1: Combinations for each sampling area and the number of census blocks groups for each unique combination of percent water and terrain roughness (U=Urban, R=Rural).

21=water percent medium, roughness small

22=water percent medium, roughness medium

23=water percent medium, roughness large

31=water percent large, roughness small

32=water percent large, roughness medium

33=water percent large, roughness large

Proposed avian and insect sampling approach

The proposed sampling approach will not only provide further information on the transmission cycle of WNV in the urban setting, but will also provide unique information regarding transmission cycles in surrounding rural settings.

The four Mississippi regions that are representative of a variety of ecological and social conditions will be sampled within both urban and rural settings. Each region will be subdivided into three sampling sites per setting. (Site 1: urban-3 sites, rural-3 sites, Site 2: urban-3 sites, rural-3 sites, etc.). Data, related to WNV transmission will be collected at each of these sites for two consecutive years on bird populations, insects, humans and the environment.

Results

Four areas of the state were chosen for sampling that broadly characterizes the range of significant ecological, hydrologic, and social variation across the state, but all four areas did not contain all 9 potential combinations. Table 1 lists the combinations for each area and the number of census blocks groups for each unique combination of percent water and terrain roughness. It is evident from this table that every unique combination of percent water and terrain roughness is not represented for each sample site. For some study sites the best case sampling scenario is developed by choosing the census block group polygons that have the most similar landscape properties possible.

Figure 2 shows the sample frame developed for Site 1. This basic sample design illustrated for Site 1 was also applied across the other study areas that included Sites 2, 3 and 4. The Site 1 urban and rural block group polygons do not both contain index value 11. In fact, there are no index values of 11 in the rural polygons and no index values of 12 in the urban polygons. Consequently, it can be seen in Figure 2 that for Site 1, the 'closest' match is urban index value = 11 and rural index value = 12. Each site in the study area contains 'special' cases where the best case sampling scenario is developed by choosing the polygons that have the most similar landscape properties possible.

Discussion and Conclusions

This paper focuses on development of a sampling scheme that minimizes redundancy for testing for WNV in bird and mosquito populations in both urban and rural environments using as an example four sites in Mississippi. The sampling technique was designed to enhance collection of information that is associated with paired landscape census block group polygons in both urban and rural environments that have similar landscape properties but vary by population density. Landform characteristics for water and terrain roughness were measured respectively, by percent water per area and the rate of slope change, the second derivative of elevation. The category thresholds for these two variables were determined through a heuristic approach designed to maximize the differences between water-roughness interactions while minimizing over-specification effects that resulted in categorization thresholds that contained very few 'paired' parcels with similar landscape properties. Nevertheless, the water-roughness interaction thresholds that were chosen on this basis still necessitated some compromises associated with choosing the 'closest' pair of census block group polygons with similar landscape properties. There are other important landscape and population dynamics variables that impact host-vector relationships critical for disease transmission [2,4]. GIS layers of landscape features including vegetation, soil texture, and other derived variables including stream density and proximity to water may be useful enhancements to the sample design. However, with each added variable the number of possible combinations for the variable interactions increases exponentially. The authors believe that the implementation of this sample design using two interactive continuous variables that measure relatively static landscape properties and a third categorical variable, population density, provides a logical first step to answer the question of whether WNV transmission dynamics are influenced by population density. The authors acknowledge that there are limitations associated with zonal-based analyses. When measurements are summarized by zone it is assumed that measurements are constant across the zone and the potential for sudden changes at the boundaries is enhanced [12]. These limitations are considered less problematic than the subjective decision process associated with bandwidth selection when employing a kernel density analysis in GIS.

It is hoped that implementation of this sample design for testing birds and mosquitos may ultimately help reveal whether WNV disease repository originates in urban environments and spreads to rural areas or vice versa. If differences are observed between urban and rural environments, such information could help to define the relative roles of bird community structure and vector communities important for explaining the distribution of reported WNV human cases. In summary, this proposed sampling strategy is based on sound principles of geospatial analysis that employ spatial autocorrelation as a mechanism for aggregating and dividing landscapes into similar units that [15], if not accounted for, result in sample biases that have the potential to dramatically increase the amount of un-partitionable variance associated with sampling for epidemiological purposes.

References

1. Silverberg LI, Demuth F (2001) Western Nile Virus: a new public health concern. The West Virginia Medical Journal 97: 157-159.
2. Cooke WH, Grala K, Wallis RC (2006) Avian GIS models signal human risk for West Nile virus in Mississippi. International Journal of Health Geographics 5: 36.
3. Mississippi Department of Health (2012) WNV Incidence Case and Mortality Rates by County in Mississippi, Mississippi.
4. LaBeaud AD, Gorman AM, Koonce J, Kippes C, McLeod J, et al. (2008) Rapid GIS-based profiling of West Nile virus transmission: defining environmental factors associated with an urban-suburban outbreak in Northeast Ohio, USA. Geospat Health 2: 215-225.
5. Liu H, Weng Q, Gaines D (2011) Geographic incidence of human West Nile virus in northern Virginia, USA, in relation to incidence in birds and variations in urban environment. Sci Total Environ 20: 4235-4241.
6. Ruiz M, Walker E, Foster E, Haramis L. Kitron U (2007) Association of West Nile virus illness and urban landscapes in Chicago and Detroit. Int J Health Geograph 6: 10-21.
7. Kronenwetter-Koepel TA, Meece JK, Miller CA, Reed KD (2005) Surveillance of above- and below-ground mosquito breeding habitats in a rural midwestern community: baseline data for larvicidal control measures against West Nile Virus vectors. Clinical Medicine & Research 3: 3-12.
8. PA (2001) Statistical Methods for Geography. Sage Publications, London, UK 232.
9. Daniel J (2012) Sampling Essentials: Practical Guidelines for Making Sampling Choices. SAGE Publications, Inc., UK.
10. http://www.census.gov/population/censusdata/urdef.txt
11. http://resources.arcgis.com/en/help/main/10.1/index.html#//009z000000w7000000
12. de Smith MJ, Goodchild MF, Longley PA (2007) Geospatial Analysis: a Comprehensive Guide to Principles, Techniques and Software Tools. Leicester, UK: Troubador.
13. Berry JK (1993) Beyond Mapping: Concepts, Algorithms and Issues in GIS. Wiley and Sons Publishers, USA.
14. Tarolli P, Sofia G, Fontana GD (2010) Geomorphic features extraction from high-resolution topography: landslide crowns and bank erosion. Nat Hazards 61: 65-83.
15. O'Sullivan D, Unwin DJ (2010) Geographic Information Analysis and Spatial Data, in Geographic Information Analysis, Second Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Citation: Cooke III WH, Grala K, Brooks CP, Pote LMW (2013) Development of a Sampling Scheme for West Nile Virus Testing in Urban and Rural Environments. Air Water Borne Diseases 2: 111. doi: 10.4172/2167-7719.1000111